Proc. Natl. Acad. Sci. USA
Vol. 83, pp. 1276–1280, March 1986
Biochemistry

A mammalian high mobility group protein recognizes any stretch of six A·T base pairs in duplex DNA

(α-protein/DNase I footprinting/netropsin/minor groove recognition of (A+T) DNA)

MARK J. SOLOMON, FRANÇOIS STRAUSS*, AND ALEXANDER VARSHAVSKY

Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139

Communicated by John M. Buchanan, October 23, 1985

ABSTRACT **α-Protein is a high mobility group protein originally purified from African green monkey cells based on its affinity for the 172-base-pair repeat of monkey α-satellite DNA. We have used DNase I footprinting to identify 50 α-protein binding sites on simian virus 40 DNA and thereby to determine the DNA binding specificity of this mammalian nuclear protein. α-Protein binds with approximately equal affinity to any run of six or more A·T base pairs in duplex DNA, to many, if not all, runs of five A·T base pairs, and to a small number of other sequences within otherwise (A+T)-rich regions. Unlike well characterized sequence-specific DNA binding proteins such as bacterial repressors, α-protein makes extensive contacts within the minor groove of B-DNA. These and related findings indicate that, rather than binding to a few specific DNA sequences, α-protein recognizes a configuration of the minor groove characteristic of short runs of A·T base pairs. We discuss possible functions of α-protein and the similarities in DNA recognition by α-protein and the antibiotic netropsin.**

Previous studies from this laboratory have addressed the existence and properties of DNA sequence-specific nucleosome-binding proteins (1–3). In particular, we searched for a protein specific for the α-satellite DNA (α-DNA) of the African green monkey. Using the "band-competition" assay, a generally applicable electrophoretic assay for specific DNA-binding proteins in crude extracts, we purified an abundant nuclear protein from green monkey CV-1 cells that preferentially bound to α-DNA (1). The solubility properties, amino acid composition, and primary structure of this ≈10 kDa protein (tentatively called α-protein) operationally classified it as a high mobility group (HMG) protein (4–6), distinct from the other major HMG proteins, HMG 1, -2, -14, and -17 (J. McCartney, F.S., M.J.S., J. Smart, and A.V., unpublished data). The preferred α-nucleosome frame detected in isolated chromatin (7, 8) is precisely bordered by α-protein binding sites (GATATTT) on α-DNA, suggesting that α-protein might function as a nucleosome-positioning or phasing protein (1).

To address the binding specificity of α-protein in more detail, we mapped α-protein binding sites on simian virus 40 (SV40) DNA. α-Protein binds with approximately equal affinity not only to the GATATTT sequences in SV40 DNA but also to >50 other sites in the ≈2.4 kilobase pairs (kbp) that we have examined by DNase I footprinting. Thus, rather than recognizing a few specific nucleotide sequences, α-protein recognizes an aspect of B-DNA conformations, most likely a configuration of the minor groove, that is characteristic of short runs of A·T base pairs.

These and other properties of α-protein set it apart from the more extensively studied prokaryotic and eukaryotic sequence-specific DNA binding proteins, whose characteristic features include the predominance of major groove interactions and little or no sequence degeneracy in DNA recognition (9).

MATERIALS AND METHODS

DNase I Footprinting. DNA fragments end-labeled with ^{32}P using either polynucleotide kinase or Klenow DNA polymerase (Bethesda Research Laboratories) were incubated in 25 μl of 70 mM NaCl/5 mM $MgCl_2$/1 mM Na EDTA/10 mM 2-mercaptoethanol/0.1% Triton X-100/4% (vol/vol) glycerol/10 mM Na Hepes, pH 7.5, for 10 min at ≈20°C with the amounts of purified α-protein (1) indicated in the figure legends. DNase I footprinting was carried out as described (1).

Analysis of α-Protein–DNA Complexes on Low Ionic Strength Gels. Purified α-protein and an end-labeled DNA fragment were incubated together for 10 min at ≈20°C in the footprinting buffer lacking $MgCl_2$, followed by electrophoresis at 4°C in a low ionic strength 5% polyacrylamide gel (1).

Interference with α-Protein Binding via Chemical Modification of DNA. The 92-bp *Dde* I/*Hin*dIII fragment of the α-DNA repeat (see Fig. 1) was 3'-end-labeled with Klenow polymerase at the *Hin*dIII end. Methylation of DNA with dimethyl sulfate (Fluka) and ethylation by ethylnitrosourea (Sigma) were performed as described (10, 11). The modified DNAs were purified by polyacrylamide gel electrophoresis. For cleavage at methylated purines, the DNA was incubated at 90°C for 15 min (pH 7.5) and then at 90°C for 30 min in 0.1 M NaOH, precipitated, and thereafter analyzed by electrophoresis on an 8% polyacrylamide sequencing gel (11). Ethylated phosphates were cleaved by incubating DNA at 90°C for 30 min in 0.15 M NaOH followed by analysis on a sequencing gel (11).

RESULTS

α-Protein Binding Sites on α-DNA. We have footprinted α-protein binding to each of the two strands of α-DNA to better define the boundaries of α-protein binding sites within the 172-bp α-DNA repeat (Fig. 1). The three protected regions are marked I–III and the extent of protection is indicated above the α-DNA sequence in Fig. 1*C*. Two features of the α-protein footprints are clear from the summary in Fig. 1*C*. First, the protected regions on the two strands are shifted 3–4 bp relative to each other because of the stagger inherent in DNase I cutting of double-stranded DNA (12). Second, when footprints of the two strands are viewed together, all three sites are seen to contain a stretch of 6 A·T base pairs (Fig. 1*C*). Site III is apparently a

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: SV40, simian virus 40; bp, base pair(s); HMG, high mobility group.

*Present address: Institut Jacques Monod, 75251 Paris 05, France.

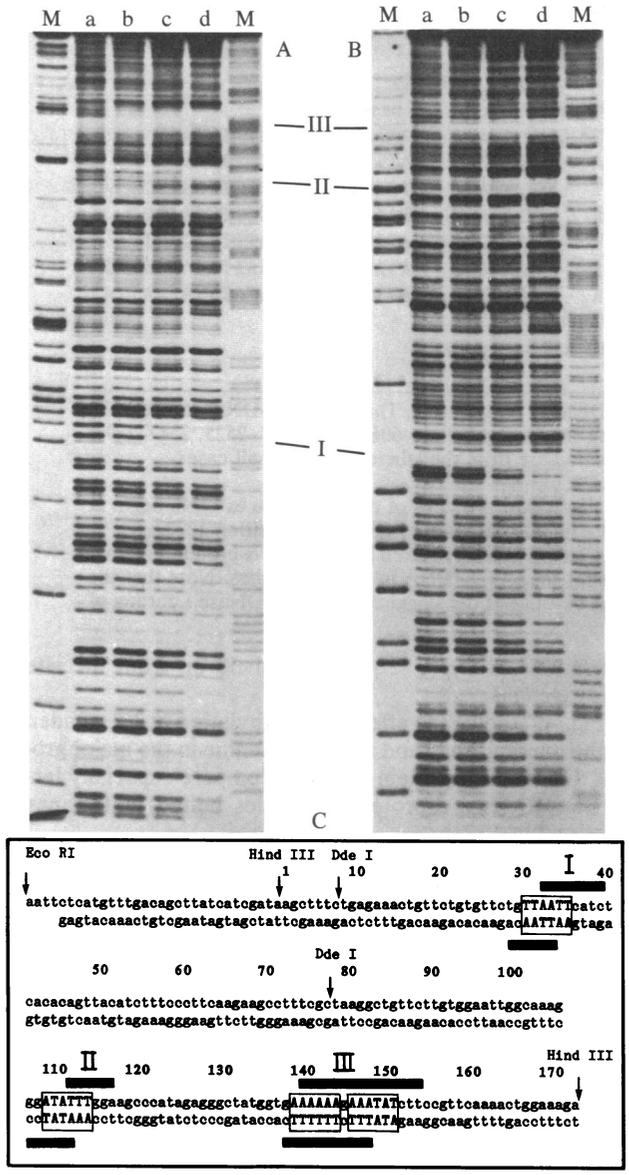

FIG. 1. α-Protein protects three sites within the 172-bp α-DNA repeat from cleavage by DNase I. (*A* and *B*) A 227-bp *Eco*RI/*Hha* I fragment from pFS522 (1) containing a single *Hin*dIII-produced α-DNA repeat was either 5'- or 3'-end-labeled at the *Eco*RI site of the vector (33). Four nanograms of DNA was digested with DNase I after incubation with 0 ng (lane a), 2.5 ng (lane b), 7.5 ng (lane c), or 23 ng (lane d) of purified α-protein. Lanes M contain size markers. The three protected sites within the α-DNA repeat are denoted I, II, and III. (*C*) Nucleotide sequence of the cloned α-DNA repeat with bars above and below indicating the extent of the protected regions on each of the two strands. The boxes denote all stretches of 6 or more contiguous A·T base pairs within the α-DNA repeat.

compound α-protein binding site containing two smaller sites separated by a single G·C base pair (Fig. 1*C*).

α-Protein Recognizes Any Stretch of Six or More A·T Base Pairs. The sequence GATATTT of sites II and III in α-DNA (Fig. 1*C*) would occur by chance on average less than once per 5 kbp. However, the 5.2-kbp SV40 genome (13) contains six such sites in two clusters of three sites each (see *Discussion* and legend to Table 1). We have analyzed α-protein binding to these sites and to a total of ≈2.4 kbp of SV40 DNA to better define the DNA binding specificity of α-protein. Table 1 lists all of the α-protein binding sites that we have identified by DNase I footprinting including those in α-DNA, SV40 DNA, and a portion of pBR322 DNA (1). All five of the GATATTT sites examined in SV40 DNA bound α-protein (Table 1). In addition, all other runs of 6 or more A·T base pairs in SV40 and pBR322 also bound α-protein. Another 20 α-protein binding sites contain 5 A·T base pairs, while two binding sites contain 4 A·T base pairs within highly (A+T)-rich stretches of DNA. Binding to 8 stretches containing runs of 5 A·T base pairs (out of 28 such stretches examined) could not be unambiguously determined due primarily to lack of DNase I cutting of naked DNA at these sites. While most if not all runs of 5 A·T base pairs are sites for α-protein binding, nearly all runs of 4 A·T base pairs are clearly not sites for α-protein binding (Table 1).

α-Protein Binding Sites Detectable by DNase I Footprinting Have Approximately Equal Affinities for α-Protein. We footprinted an SV40 restriction fragment in the presence of increasing amounts of α-protein to compare the relative affinities of seven α-protein binding sites present in this DNA fragment (Fig. 2). At least a partial protection of all seven sites is observed at a particular concentration of α-protein (lane 5) and complete protection of all sites occurs at a 3-fold higher concentration of α-protein (lane 6). Furthermore, no additional protected sites appeared upon further increases in α-protein concentration (ref. 1; data not shown). As shown in Fig. 2*C*, sites b–e and g contain 6 A·T base pairs (site c is identical to sites II and III found in α-DNA), while sites a and f contain only 5 A·T base pairs. Despite this remarkable diversity in sequence, all seven sites have approximately the same affinity for α-protein.

α-Protein Preferentially Binds to Double-Stranded DNA. One property of (A+T)-rich stretches of duplex DNA is their relatively low melting temperature. Thus, the strong preference of α-protein for (A+T) DNA could be explained if α-protein preferentially bound to single-stranded DNA. To directly address this point, we incubated α-protein with a mixture of isolated single strands of α-DNA, double-stranded α-DNA, and increasing amounts of unlabeled double-stranded *Escherichia coli* competitor DNA (Fig. 3). Essentially no α-protein is bound to single-stranded α-DNA at an *E. coli* competitor DNA concentration at which most of the duplex α-DNA remains complexed with at least one molecule of α-protein (lanes d and k). We estimate that the preference of α-protein for duplex α-DNA relative to single-stranded α-DNA is at least 5-fold and may actually be higher, especially if secondary structures within the single strands of α-DNA reform α-protein binding sites. This preference of α-protein for double-stranded DNA makes untenable the model of α-protein–DNA recognition via local melting of the double helix, as that would require a strong preference for single-stranded DNA.

α-Protein Makes Phosphate and Minor Groove Contacts with Duplex DNA. To probe the contacts between α-protein and its binding sites on DNA we modified the α-DNA prior to α-protein binding with either dimethyl sulfate or ethylnitrosourea (10). α-Protein was added to a mixture of a labeled restriction fragment of modified α-DNA containing binding sites II and III (see Fig. 1*C*) and decreasing amounts of unlabeled competitor *E. coli* DNA. Complexes of α-protein and α-DNA were separated by low ionic strength gel electrophoresis (1, 14–16). The DNAs of electrophoretic bands containing 0, 1, or 2 α-protein molecules per molecule of α-DNA (Fig. 4*A*) were eluted, cleaved at the modified positions, and analyzed on sequencing gels (Fig. 4 *B* and *C*). Ethylation at any of at least four phosphates within the sequence ATATTT (site II) strongly interferes with binding of α-protein (compare especially lane II with lane 0 in Fig. 4*B*). Interestingly, an ethylation-induced DNA modification 16 bases from the center of this α-protein binding site also strongly interferes with α-protein binding (Fig. 4*B*, arrowhead). Our interpretation is that this modification influences the conformation of B-DNA within the binding site. Poten-

Table 1. α-Protein binding sites

Binding sites on α-DNA							
TTAATT		ATATTT		AAAAAA		AAATAT	
Binding sites on SV40 and pBR322 containing five or more A·T base pairs							
AAAAT	AATTT	TTTAA	AAAATA	TTAAATT*	AAATAAA	AAAATAT†	ATATTTAA†
AATTA	ATTTT	TTTAT	ATTATA	ATATAAA	AAATAAA	ATTTTTTT	TTTAAAAAA
AAAAA	AAAAT	AATTT	TTAAAT	AAAAAAT	TTATAAT	TAATTAAT	TTTTTAATTT
ATTTT	ATTTT	AAAAT	AAATTT	TTTTAAA	ATAAAAT	TAAAAAAT	TTTTTAAATAT
ATAAT	TTTTA	ATAATA	TTTATT	AATATTT	TAATTAA	ATTTTAAT	ATTTTATATTTA
TTATT	TTTTT	TTTTTT	TTTTTA	TTAATAA	AAAAAAT	AAATTATA	AAATAAAATATAT†
AAAAT	TTTAT	AATTAT	TTTAAT*	TAAAATA	AAATTTT	ATTTTTTT	ATATTTAAAATTA†
TTTAT	TTAAA	AATATT	AAATAT†	TTTTTTT			
Binding sites on SV40 with four A·T base pairs‡							
gccTATAcAAATcTAc		ggAAAcTAAAcAAgTg					

Shown are all of the α-protein binding sites we have detected by DNase I footprinting. The sites on α-DNA are from the data in Fig. 1. We examined positions 30–100 of pBR322 for binding sites and positions 606–1232, 1913–2525, 2547–2792, 4017–4158, 4392–4727, and 4756–5231 of SV40 (13). The sites indicated are from the labeled strand in all cases.

*Derived from footprint analysis of α-protein binding sites in pBR322.

†Identical to sites II and III (GATATTT/AAATATC) on α-DNA. The six occurrences of this sequence in SV40 begin at positions 952, 1118, 1277, 2018, 2163, and 2358 (13). We have not examined binding to the site at position 1277.

‡Underlining indicates the extent of the regions on the labeled DNA strand protected from DNase I by α-protein binding. Guanines and cytosines are shown in lower case letters to emphasize the (A+T) DNA in these stretches. Note that the actual binding sites are ≈2 bases to the left of the protected regions because of the stagger inherent in DNase I cutting of B-DNA (see Fig. 1*C* and ref. 12).

tially analogous long-range effects are seen with bleomycin-induced cleavage of duplex DNA, which can be modulated by alterations in DNA sequence over 50 bp from the site of cleavage (17).

Fig. 4*C* shows the effects of methylation at N3 of adenine in the minor groove and at N7 of guanine in the major groove (10, 11, 18) on α-protein binding. Compare especially lane II in Fig. 4*C* (two α-protein molecules bound) with lanes M

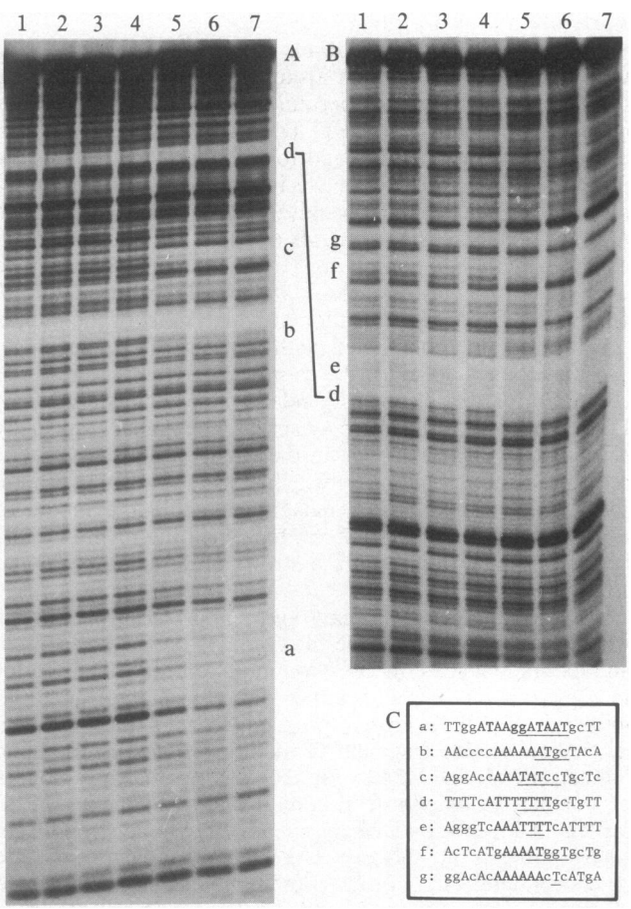

FIG. 2. Different α-protein binding sites have similar affinities for α-protein. One nanogram of a 478-bp *Mbo* I/*Aha* III fragment of SV40 DNA 3′-end-labeled at the *Mbo* I site was digested with DNase I in the presence of 0 ng (lane 1), 0.2 ng (lane 2), 0.6 ng (lane 3), 1.9 ng (lane 4), 5.5 ng (lane 5), 17 ng (lane 6), or 50 ng (lane 7) of purified α-protein. The seven protected regions are indicated by the letters a–g. In *B*, the same samples as in *A* were electrophoresed for a 3-fold longer time to resolve sites e–g. The nucleotide sequences of the labeled strand encompassing sites a–g are shown in *C* with the protected regions underlined.

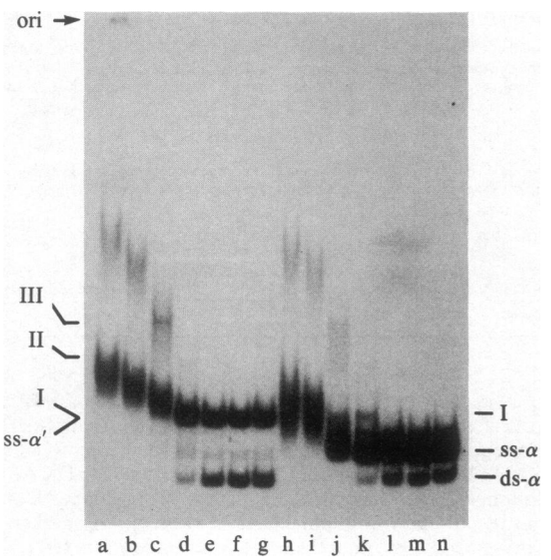

FIG. 3. α-Protein binds more tightly to double-stranded than to single-stranded DNA. The two end-labeled strands of a single 172-bp α-DNA repeat (denoted ss-α and ss-α′) were electrophoretically separated (11). Each of the purified single strands was mixed with the end-labeled double-stranded α-DNA fragment (ds-α) (0.1 ng total of α-DNA), 2 ng of purified α-protein, and binding buffer in the presence of 0 ng (lanes a and h), 1.6 ng (lanes b and i), 6 ng (lanes c and j), 25 ng (lanes d and k), 100 ng (lanes e and l), or 400 ng (lanes f and m) of double-stranded *E. coli* competitor DNA (≈1 kbp). Lanes a–g contained single-stranded α-DNA; lanes h–m contained single-stranded α′-DNA. Lanes g and n contained neither α-protein nor *E. coli* DNA. Binding of α-protein to DNA was assayed by low ionic strength electrophoresis (see *Materials and Methods* and refs. 1 and 14). I, II, and III denote positions of complexes of one, two, and three α-protein molecules with the double-stranded α-DNA fragment. ori, Origin. Weak bands in the naked DNA lanes (g and n) are probably due to minor alternative conformations of single-stranded DNA fragments. Naked single-stranded α′-DNA and complex I comigrate in this electrophoretic system.

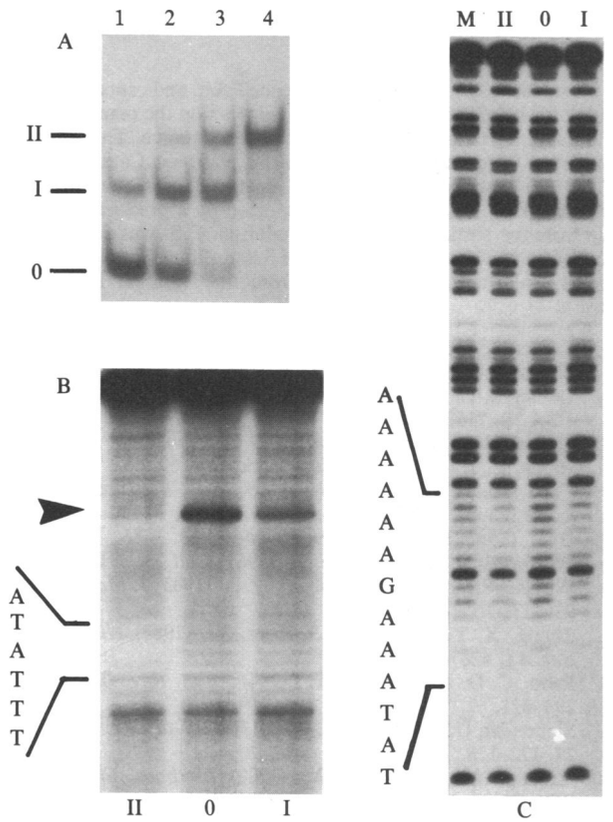

FIG. 4. α-Protein makes minor groove and phosphate contacts with the DNA. An end-labeled 92-bp *Hin*dIII/*Dde* I fragment of α-DNA (see Fig. 1*C*) was methylated with dimethyl sulfate or ethylated with ethylnitrosourea. (*A*) Approximately 15 ng of chemically modified DNA was incubated in 75 μl of binding buffer with 1.5 ng of purified α-protein and 150 ng (lane 1), 75 ng (lane 2), 36 ng (lane 3), or 18 ng (lane 4) of double-stranded *E. coli* competitor DNA (≈1 kbp). Complexes containing two (II), one (I), or no (0) molecules of α-protein were resolved by low ionic strength electrophoresis. The example shown in *A* represents an experiment with methylated α-DNA; no differences in electrophoretic mobility were observed between methylated, ethylated, and untreated DNA. (*B* and *C*) DNAs of electrophoretic bands 0, I, and II in lane 3 were eluted, selectively cleaved at the sites of ethylation (*B*) or methylation (*C*), and electrophoresed on 8% polyacrylamide sequencing gels. Numerals 0, I, and II in *B* and *C* denote the numbers of α-protein molecules bound in the complex. Arrowhead in *B* indicates the position of a preferential site of ethylation outside of the α-protein binding sites that nonetheless interferes with α-protein binding. Lane M shows the cleavage pattern of methylated DNA before fractionation as in *A*. Nucleotide sequences of site II and of the bipartite site III are shown in *B* and *C*, respectively.

(control) and 0 (unbound fraction). Methylation of any of the adenines within the sequence AAAAAAGAAATAT interferes with α-protein binding. Interference with α-protein binding is not as strong as in the ethylation experiment (Fig. 4*B*), possibly because the binding of α-protein to one of the two binding sites within the compound site III (Fig. 1*C*) precludes or decreases binding to the adjacent site.

The 2-Amino Group of Guanine Prevents High-Affinity α-Protein Binding to (G+C) DNA. The above methylation interference data indicate that α-protein recognizes (A+T) DNA via minor groove interactions. As seen from the minor groove, G·C and A·T base pairs differ solely by the presence of the 2-amino group in guanine instead of the H atom in adenine. Replacement of the 2-amino group with H yields inosine (I). Thus the I·C base pair resembles G·C in the major groove and A·T in the minor groove. To address the role of this NH_2 group in α-protein–DNA recognition, we compared the relative abilities of total *E. coli* DNA, poly(dA-dT)·poly(dA-dT), poly(dG-dC)·poly(dG-dC), and poly(dI-dC)·poly(dI-dC) to compete with α-DNA for α-protein binding (Fig. 5). Although 4 ng of *E. coli* DNA eliminates most α-protein binding to α-DNA (Fig. 5*A*, lane d), ≈30 ng of poly(dG-dC)·poly(dG-dC) DNA is required for the same degree of competition (Fig. 5*C*, lanes e and f). On the other hand, poly(dA-dT)·poly(dA-dT) competes for α-protein binding ≈8-fold better than total *E. coli* DNA (compare Fig. 5*B*, lane b with Fig. 5*A*, lane c), and ≈100-fold better than poly(dG-dC)·poly(dG-dC) DNA (compare Fig. 5*B* with Fig. 5*C*). Poly(dI-dC)·poly(dI-dC) competes as well as poly(dA-dT)·poly(dA-dT), demonstrating that it is the presence of the 2-amino group of guanine in the minor groove that prevents high-affinity binding of α-protein to (G+C) DNA.

DISCUSSION

Our findings indicate that α-protein binds to a great variety of (A+T) DNA sequences because of its recognition of a specific aspect of B-DNA conformations characteristic of any run of 6 or more A·T base pairs, and also of most, if not all, runs of 5 A·T base pairs. However, acquisition of such "binding" DNA conformations by shorter stretches of (A+T) DNA appears to be strongly influenced by flanking DNA sequences.

α-Protein makes extensive minor groove contacts (see *Results*), consistent with theoretical predictions for degenerate recognition of (A+T) DNA (19). This result, together with the data on α-protein binding to synthetic DNAs (see *Results*), indicates that the structure of the minor groove underlies the sequence-degenerate recognition of (A+T) DNA by α-protein. Histones within the nucleosome also contact the DNA primarily within the minor groove (20, 21), potentially analogous to α-protein–DNA interactions. A more detailed analogy is provided by the mechanism of DNA recognition by the nonintercalative, low molecular weight antibiotic netropsin, which binds within the minor groove to

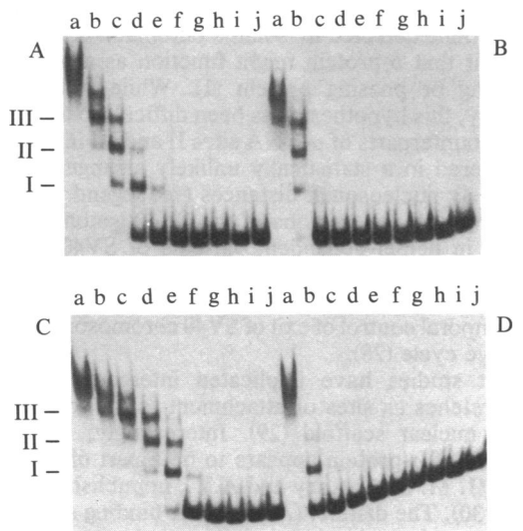

FIG. 5. G·C base pairs interfere with α-protein binding via the 2-amino group of guanine in the minor groove. Purified α-protein (0.8 ng) was added to a mixture of the end-labeled 172-bp α-DNA fragment (1 ng) and either sonicated *E. coli* DNA (*A*), poly(dA-dT)·poly(dA-dT) (*B*), poly(dG-dC)·poly(dG-dC) (*C*), or poly(dI-dC)·poly(dI-dC) (*D*). All polynucleotides were obtained from P-L Biochemicals. α-Protein–α-DNA complexes were resolved on low ionic strength polyacrylamide gels. Samples in lanes a–i contained 0, 0.25, 1, 4, 15, 60, 250, 1000, and 4000 ng of competitor DNA, respectively. No α-protein was added to samples in lane j. I, II, and III denote positions of complexes of α-DNA with one, two, and three molecules of α-protein, respectively.

clusters of at least 4 A·T base pairs (22). The DNA binding specificity of netropsin results largely not from specific hydrogen bonding but from close van der Waals contacts between C-2 hydrogens of adenine in the minor groove and CH groups of the multiple pyrrole rings of netropsin (23). These interactions are sterically prevented by the 2-amino group of guanine, accounting for the much lower affinity of netropsin for (G+C) DNA. Removal of this NH_2 group from guanine (to yield I) results in high-affinity binding of both netropsin (24) and α-protein to poly(dI-dC)·poly(dI-dC) DNA (see *Results*). Future x-ray analysis of α-protein–DNA complexes will determine whether the striking similarity of minor groove-mediated DNA recognition by α-protein and netropsin is due to a similarity of contacts seen at atomic resolution.

Unlike the other major HMG proteins, which preferentially bind single-stranded DNA *in vitro* (4, 5, 20), α-protein preferentially binds double-stranded DNA (see *Results*). Moreover, we have failed to detect any sequence specificity of HMG14 or HMG17 binding to double-stranded DNA (unpublished data) using the "band-competition" assay of the type used to detect and purify α-protein (1). It remains to be seen whether the distinct physicochemical properties that define the family of HMG proteins reflect an underlying functional similarity.

Lund *et al.* (25) have recently isolated three closely related human HMG proteins (HMG-I, HMG-Y, and HMG-M), one of which (HMG-I) is identical to α-protein. α-Protein (HMG-I), which is itself a phosphoprotein, is also phosphorylated at mitosis to yield HMG-M, and it is possible that HMG-Y is yet another phosphorylated counterpart of α-protein (25). Neither the functional significance of these multiple phosphorylations nor their effects on the DNA binding specificity of α-protein have been explored. It is also unknown whether the (A+T) DNA binding specificity of α-protein seen *in vitro* with naked DNA ligands is either retained or further restricted within chromatin *in vivo*.

The function of α-protein is not known. The demonstration that binding sites II and III within the 172-bp repeat of α-DNA are located at the boundaries of the preferred α-nucleosome phasing frame detected in isolated chromatin (1, 7, 8) led us to suggest that α-protein might function as a nucleosome-positioning or phasing protein (1). While still a distinct possibility, this hypothesis has been difficult to test directly. The six counterparts of α-DNA sites II and III in SV40 DNA are clustered in a statistically unlikely arrangement at approximately nucleosomal distances (see legend to Table 1). Interestingly, these portions of the SV40 genome are often enriched in helper-dependent variants of SV40 containing reiterated subgenomic sequences (26, 27). In addition, one of the site II/III clusters is contained within a region implicated in the temporal control of exit of SV40 chromosomes from the replicative cycle (28).

Recent studies have implicated intergenic (A+T)-rich DNA stretches as sites of attachment to the (operationally defined) nuclear scaffold (29). Interestingly, a significant proportion of α-protein appears to be a part of the nuclear scaffold (J. M. McCartney and A.V., unpublished data; see also ref. 30). The distinct (A+T) DNA binding specificity of α-protein and its high relative content in the nucleus are thus consistent, among other possibilities, with a role in nuclear scaffold–DNA interactions *in vivo*.

α-Protein is found in cultured mammalian cells ranging from human to murine (ref. 1 and unpublished data). Several higher molecular weight (A+T) DNA binding proteins have also been reported in nonmammalian species (3, 31, 32). For instance, D1, an abundant ≈55 kDa nuclear protein from *Drosophila melanogaster* (2, 3, 31) recognizes stretches of (A+T) DNA *in vitro* with a specificity similar if not identical to that of α-protein (R. Pan, F.S., and A.V., unpublished data). It remains to be seen whether the similarity of DNA binding properties underlies homologous functions for these diverse proteins.

We thank John McCartney, Lawrence Peck, and especially Daniel Finley for helpful discussions and comments on the manuscript. We also thank Barbara Doran for secretarial assistance. This work was supported by grants to A.V. from the National Cancer Institute (CA30367) and from the National Institute of General Medical Sciences (GM33401). M.S. was supported by a predoctoral fellowship from the National Science Foundation. F.S. was supported by the Centre National de la Recherche Scientifique (France) and by a Fogarty International Fellowship from the National Institutes of Health.

1. Strauss, F. & Varshavsky, A. (1984) *Cell* **37,** 889–901.
2. Levinger, L. & Varshavsky, A. (1982) *Cell* **28,** 375–385.
3. Levinger, L. & Varshavsky, A. (1982) *Proc. Natl. Acad. Sci. USA* **79,** 7152–7156.
4. Walker, J. M. (1982) in *The HMG Chromosomal Protein*, ed. Johns, E. J. (Academic, New York), pp. 69–88.
5. Cartwright, I. L., Abmayr, S. M., Fleischmann, G., Lowenhaupt, K., Elgin, S. C. R., Keene, M. A. & Howard, G. C. (1982) *CRC Crit. Rev. Biochem.* **13,** 1–86.
6. Weisbrod, S. & Weintraub, H. (1981) *Cell* **23,** 391–400.
7. Wu, K., Strauss, F. & Varshavsky, A. (1983) *J. Mol. Biol.* **170,** 93–117.
8. Zhang, X. Y., Fittler, F. & Horz, W. (1983) *Nucleic Acids Res.* **11,** 4287–4305.
9. Pabo, C. O. & Sauer, R. T. (1984) *Annu. Rev. Biochem.* **53,** 293–321.
10. Siebenlist, U. & Gilbert, W. (1980) *Proc. Natl. Acad. Sci. USA* **77,** 122–126.
11. Maxam, A. & Gilbert, W. (1980) *Methods Enzymol.* **65,** 499–525.
12. Camenini-Otero, R. D., Sollner-Webb, B., Simon, R. H., Williamson, P., Zasloff, M. & Felsenfeld, G. (1978) *Cold Spring Harbor Symp. Quant. Biol.* **42,** 57–75.
13. Tooze, J., ed. *DNA Tumor Viruses* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY), pp. 801–804.
14. Fried, M. G. & Crothers, D. M. (1984) *J. Mol. Biol.* **172,** 241–262.
15. Varshavsky, A., Bakayev, V. V. & Georgiev, G. P. (1976) *Nucleic Acids Res.* **3,** 477–492.
16. Hendrickson, W. & Schleif, R. (1985) *Proc. Natl. Acad. Sci. USA* **82,** 3129–3133.
17. Murray, V. & Martin, R. F. (1985) *Nucleic Acids Res.* **13,** 1467–1481.
18. Sakonju, S. & Brown, D. D. (1982) *Cell* **31,** 395–405.
19. Rosenberg, J. M. & Greene, P. (1982) *DNA* **1,** 117–124.
20. McGhee, J. D. & Felsenfeld, G. (1980) *Annu. Rev. Biochem.* **49,** 1115–1156.
21. Richmond, T. J., Finch, J. T., Rushton, B., Rhodes, D. & Klug, A. (1984) *Nature (London)* **311,** 532–537.
22. Van Dyke, M. W., Hertzberg, R. P. & Dervan, P. B. (1982) *Proc. Natl. Acad. Sci. USA* **79,** 5470–5474.
23. Kopka, M. L., Yoon, C., Goodsell, D., Pjura, P. & Dickerson, R. E. (1985) *Proc. Natl. Acad. Sci. USA* **82,** 1376–1380.
24. Wartell, R. M., Larson, J. E. & Wells, R. D. (1974) *J. Biol. Chem.* **240,** 6719–6731.
25. Lund, T., Holtland, J. & Laland, S. G. (1985) *FEBS Lett.* **180,** 275–279.
26. McCutchan, T., Singer, M. & Rosenberg, M. (1979) *J. Biol. Chem.* **254,** 3592–3597.
27. Sheflin, L., Celeste, A. & Woodworth-Gutai, M. (1983) *J. Biol. Chem.* **258,** 14315–14321.
28. Wang, H. T., Larsen, P. H. & Roman, A. (1985) *J. Virol.* **53,** 410–414.
29. Mirkovitch, J., Mirault, M. E. & Laemmli, U. (1984) *Cell* **39,** 223–232.
30. Sahyoun, N., LeVine, H., Bronson, D. & Cuatrecasas, P. (1984) *J. Biol. Chem.* **259,** 9341–9344.
31. Rodriguez-Alfageme, C. R., Rudkin, G. T. & Cohen, L. H. (1980) *Chromosoma* **78,** 1–31.
32. Garreau, H. & Williams, J. G. (1983) *Nucleic Acids Res.* **11,** 8473–8484.
33. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY).